PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : H04L 12/54

A1

(11) International Publication Number: WO 00/67433

(43) International Publication Date: 9 November 2000 (09.11.00)

(21) International Application Number: PCT/US00/10631

(22) International Filing Date: 19 April 2000 (19.04.00)

(30) Priority Data:
09/303,861 3 May 1999 (03.05.99) US

(71) Applicant: MOTOROLA INC. [US/US]; 1303 East Algonquin Road, Schaumburg, IL 60196 (US).

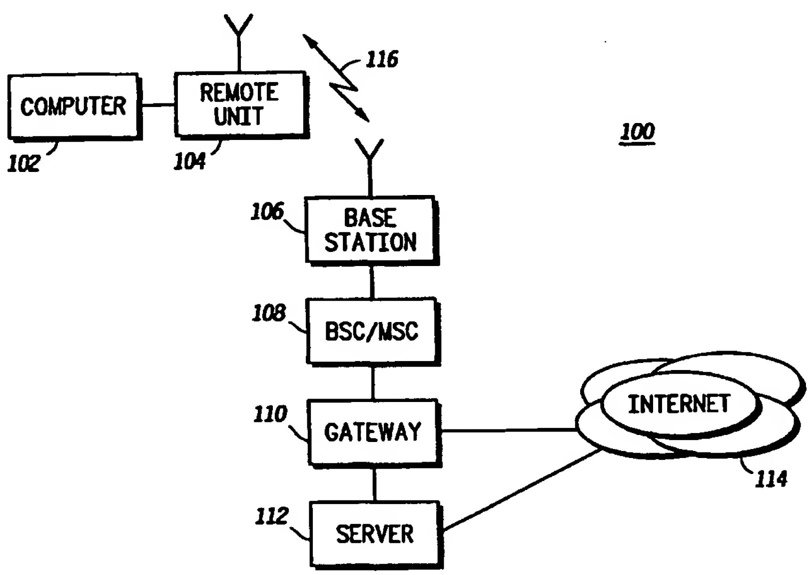
(72) Inventors: RAMANNA, Shreesha; 407 Happfield Drive, #201, Arlington Heights, IL 60004 (US). PAZHYANNUR, Rajesh, S.; 1348 Geneva Drive, #1B, Palatine, IL 60067 (US). JACOB, Kurian, P.; 4988 Thornbark Drive, Barrington, IL 60010 (US). JAYAPALAN, Jay, P.; 1215 Brandywyn Lane, Buffalo Grove, IL 60089 (US).

(74) Agents: HAAS, Kenneth, A. et al.; Motorola Inc., Intellectual Property Dept., 1303 East Algonquin Road, Schaumburg, IL 60196 (US).

(81) Designated States: BR, JP, KR, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.

(54) Title: METHOD AND APPARATUS FOR OPTIMIZING A BUFFER



(57) Abstract

When data transmission is desired, a TCP/IP session is set up between a server (112) and a remote unit (104). As part of a negotiation between the server (112) and remote unit (104), a Maximum Segment Size (MSS) is established for data transmission. Both a gateway (110) and a BSC/MSC (108) vary internal buffer sizes based on the MSS. More particularly, the gateway (110) and the BSC/MSC (108) contain internal buffers that dynamically vary in size based on the negotiated MSS. Data transmission then takes place between the server (112) and the remote unit (104) utilizing the buffers.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

METHOD AND APPARATUS FOR OPTIMIZING A BUFFER

Field of the Invention

The present invention relates generally to optimizing a buffer and, in particular, to optimizing a buffer during a Transmission Control Protocol/Internet Protocol (TCP/IP) session.

Background of the Invention

During a file transfer using a Transmission Control Protocol/Internet Protocol (TCP/IP) session, data is usually stored in buffers in order to prevent packet drops. The buffers are usually placed in network equipment having higher data input than output. Data traffic using TCP/IP over low-speed cellular networks is very sensitive to buffer size. The sensitivity manifests itself through packet drops, timeouts, and eventually, reduced throughput. Therefore, a need exists for a method and apparatus for a optimizing a buffer that reduces packet drops, timeouts, and improves packet throughput.

Brief Description of the Drawings

FIG. 1 is a block diagram of a communication system in accordance with the preferred embodiment of the present invention.

FIG. 2 illustrates TCP/IP throughput based on buffer size for Maximum Segment Sizes (MSSs) of 500, 1000, and 1500 bytes.

FIG. 3 is a block diagram of a buffer in accordance with the preferred embodiment of the present invention.

FIG. 4 is a flow chart illustrating operation of the buffer of FIG. 3 in accordance with the preferred embodiment of the present invention.

FIG. 5 is a flow chart illustrating operation of the buffer of FIG. 3 in accordance with an alternate embodiment of the present invention.

Detailed Description of the Drawings

To address the above-discussed need, a method and apparatus for optimizing a buffer is described herein. In accordance with the preferred embodiment of the present invention, when data transmission is desired, a TCP/IP session is set up between a server and a remote unit. As part of a negotiation between the server and remote unit, a Maximum Segment Size (MSS) is established for data transmission. Both a gateway and a BSC/MSC vary internal buffer sizes based on the MSS. More particularly, the gateway and the BSC/MSC contain internal buffers that dynamically vary in size based on the negotiated MSS. Data transmission then takes place between the server and the remote unit utilizing the dynamically-adjusted buffers.

The present invention encompasses a method for optimizing a buffer, the method comprising the steps of determining a segment size for a data transmission, and varying a buffer size based on the determined segment size. In the preferred embodiment of the present

invention the step of determining the segment size comprises the step of determining a Maximum Segment Size (MSS) for a TCP/IP session.

The present invention additionally encompasses a method for optimizing a buffer, the method comprising the steps of receiving a packet for a first data-transmission session and receiving a packet for a second data-transmission session. A determination is made to whether a number of packets buffered for the first data-transmission session exceeds a threshold and to whether a number of packets buffered for the second data-transmission session exceeds the threshold. The packet for the first data-transmission session is dropped if the number of packets buffered for the first data-transmission session exceeds the threshold and the packet for the second data-transmission session is dropped if the number of packets buffered for the second data-transmission session exceeds the threshold.

Finally, the present invention encompasses an apparatus comprising a variable length buffer having data and a segment size as an input and outputting buffered data. In the preferred embodiment of the present invention the buffered data was previously stored within the variable length buffer only if the variable length buffer contained an amount of data less than a threshold based on the segment size.

Turning now to the drawings, where like numerals designate like components, FIG. 1 is a block diagram of communication system 100 in accordance with the preferred embodiment of the present invention. In the preferred embodiment of the present invention, communication system 100 utilizes a Code Division Multiple Access (CDMA) system protocol as described in Cellular System Remote unit-Base Station Compatibility

Standard of the Electronic Industry Association/Telecommunications Industry Association Interim Standard 95 (TIA/EIA/IS-95B), which is incorporated by reference herein. (EIA/TIA can be contacted at 2001 Pennsylvania Ave. NW Washington DC 20006). In alternate embodiments communication system 100 may utilize other analog or digital cellular communication system protocols such as, but not limited to, the Narrowband Advanced Mobile Phone Service (NAMPS) protocol, the Advanced Mobile Phone Service (AMPS) protocol, the Global System for Mobile Communications (GSM) protocol, the Personal Digital Cellular (PDC) protocol, or the United States Digital Cellular (USDC) protocol.

Communication system 100 includes remote unit 104, base station 106, Base Station Controller/Mobile Switching Center (BSC/MSC) 108, gateway 110, and server 112. It is contemplated that network elements within communication system 100 are configured in well known manners with processors, memories, instruction sets, and the like, which function in any suitable manner to perform the function set forth herein. As shown, computer 102 and server 112 are coupled to remote unit 104 and internet 114, respectively. In the preferred embodiment of the present invention, base station 106 is preferably a Motorola base station and BSC/MSC 108 is preferably a Motorola BSC and/or MSC component. Additionally, communication between remote unit 104 and base station 106 takes place utilizing over-the-air communication signal 116.

Operation of communication system 100 in accordance with the preferred embodiment of the present invention occurs as follows: When data transmission between computer 102 and internet 114 is desired, a

data transmission session is set up between server 112 and remote unit 104. In the preferred embodiment of the present invention the data transmission session is a Transmission Control Protocol/Internet Protocol (TCP/IP) session, however, in alternate embodiments of the present invention, other transmission control protocols may be utilized. As part of a negotiation between server 112 and remote unit 104, an MSS is established for data transmission. In particular the MSS is a maximum TCP packet size allowed for the negotiated session. The MSS is based on the largest frame size that a link layer protocol can handle, the processing requirements of the two end points, and on the reliability of the network. Although the MSS is fixed for the duration of the TCP/IP session, the MSS varies between sessions based on negotiations between the sender (server 112) and the receiver (computer 102), typically at the start of the session. It should be noted that the MSS is negotiated strictly between the end points and intermediate elements like BSC/MSC cannot modify it.

In the preferred embodiment of the present invention both gateway 110 and BSC/MSC 108 determine the MSS for the purpose of buffer management and vary internal buffer sizes based on the MSS. More particularly, gateway 110 and BSC/MSC 108 contain internal buffers (shown in FIG. 2) that dynamically vary in size based on the negotiated MSS. In the preferred embodiment of the present invention buffers in both gateway 110 and BSC/MSC 108 are adjusted to three times the MSS. When multiple TCP/IP sessions are conducted by a single remote unit, gateway 110 and BSC/MSC 108 adjust internal buffers to three times the largest negotiated MSS. Data transmission then takes

place between server 112 and remote unit 104 utilizing the buffers.

FIG. 2 illustrates TCP/IP throughput based on buffer size for MSSs of 500, 1000, and 1500 bytes. The buffer size in x*MSS are depicted on the x axis (where x is an integer). Thus, a 500 byte MSS with a range of 2-6 on the x-axis implies a buffer of 1 to 3 Kbytes, while for a 1500 byte MSS the range of 2-6 implies 3-9 Kbytes. As is evident the benefits of the buffering mechanism described are two fold. Firstly, buffering in accordance with the preferred embodiment reduces the total buffer size required at the BSC/MSC and gateway, and secondly, buffering in accordance with the preferred embodiment of the present invention prevents the reduction in the TCP throughput due to buffers greater or lesser than three times MSS.

FIG. 3 is a block diagram of variable-length buffer 300 in accordance with the preferred embodiment of the present invention. As discussed above, buffer 300 resides internal to gateway 110 and BSC/MSC 108, however, in alternate embodiments of the present invention, buffer 300 may reside in other network elements (e.g., routers, intermediate servers and network elements). As is evident, buffer 300 has a data input path and a data output path, as well as an MSS input. Operation of buffer 300 in accordance with the preferred embodiment of the present invention occurs as described in FIG. 4.

The logic flow begins at step 400 where it is determined if a data transmission between remote unit 104 and server 112 needs to take place. If at step 400, it is determined that data transmission does not need to take place, the logic flow returns to step 400, otherwise the logic flow continues to step 401 where a

data transmission session is set up between server 112 and remote unit 104. In the preferred embodiment of the present invention the data transmission session is a TCP/IP session, however in alternate embodiments of the preferred embodiment, the data transmission session may utilize other transmission protocols such as ISO transport protocols and any modified TCP for wireless applications. At step 403 buffer 300 determines the current MSS. As discussed above, the MSS is a negotiated maximum segment size and during file transfer, the TCP segments are usually the same length as the MSS. In response to the MSS, buffer 300 adjusts its size to three times the MSS (step 405). At step 407 data transmission proceeds at optimum throughput utilizing the adjusted buffer. Finally at step 408 TCP/IP connection ends. The logic flow then returns to step 400 until further data transmission needs to take place.

FIG. 5 is a flow chart illustrating operation of the buffer of FIG. 3 for two TCP/IP connections in accordance with an alternate embodiment of the present invention. In the alternate embodiment of the present invention, multiple TCP/IP connections for multiple users are supported using a single buffer element. For each TCP connection, the number of packets buffered is monitored separately, and when the limit exceeds a threshold (e.g., 3*MSS), new packets are dropped. This implementation is equivalent to maintaining a buffer with size 3*MSS for each user, yet all users share a common buffer.

The logic flow begins at step 501 where buffer 300 receives a packet for a first data-transmission session (e.g., TCP/IP session) and determines a threshold (MSS). At step 503, buffer 300 receives a packet for a

second data-transmission session (e.g., TCP/IP session). Next, at step 505 buffer 300 determines if a number of packets buffered for the first connection exceeds a threshold. If, at step 505, buffer 300 determines that the number of packets buffered for the first TCP/IP session exceeds the threshold, then the logic flow continues to step 507 where the packet is dropped (i.e., not buffered), and the logic flow continues to step 511. If, at step 505, buffer 300 determines that the number of packets buffered for the first TCP/IP session does not exceed the threshold, then the logic flow continues to step 509 where the packet is buffered in buffer 300. The logic flow continues to step 511.

At step 511 buffer 300 determines if a number of packets buffered for the second connection exceeds the threshold. If, at step 511, buffer 300 determines that the number of packets buffered for the second TCP/IP session exceeds the threshold, then the logic flow continues to step 513 where the packet is dropped (i.e., not buffered), and the logic flow returns to step 501. If, at step 511, buffer 300 determines that the number of packets buffered for the second TCP/IP session does not exceed the threshold, then the logic flow continues to step 515 where the packet is buffered in buffer 300. The logic flow returns to step 501.

The descriptions of the invention, the specific details, and the drawings mentioned above, are not meant to limit the scope of the present invention. For example, in the preferred embodiment of the present invention the MSS is determined by monitoring the messaging that takes place during TCP/IP negotiation, however, in alternate embodiments of the present invention, the MSS may be determined in other ways.

For example, the MSS could be determined at the BSC/MSC based on the PPP frame size and IP packet type (TCP or UDP). In this situation the MSS information does not have to be signaled across the BSC/MSC and gateway. Alternatively, the gateway can determine the MSS and notify the BS/MSC. Additionally it should be noted that a buffer size may be adjusted in several ways. For example, physical memory may be allocated for storage, where the memory allocated is changed based on segment size. Alternatively, a static block of memory may be allocated for buffering, and packets may be withheld from storage if the current amount of data stored on the buffer exceeds a threshold. It is the intent of the inventors that various modifications can be made to the present invention without varying from the spirit and scope of the invention, and it is intended that all such modifications come within the scope of the following claims and their equivalents.

Claims

1. A method for optimizing a buffer, the method comprising the steps of:

determining a segment size for a data transmission to produce a determined segment size; and

varying a buffer size based on the determined segment size.

2. The method of claim 1 wherein the step of determining the segment size comprises the step of determining a Maximum Segment Size (MSS).

3. The method of claim 2 wherein the step of determining the MSS comprises the step of determining the MSS for a Transmission Control Protocol/Internet Protocol (TCP/IP) session.

4. The method of claim 2 wherein the step of varying the buffer size based on the determined segment size comprises the step of varying the buffer size to approximately 3*MSS.

5. The method of claim 1 wherein the step of determining a segment size for a data transmission comprises the step of determining a segment size for a Transmission Control Protocol/Internet Protocol (TCP/IP) session.

6. An apparatus comprising:

a variable length buffer having data and a segment size as an input and outputting buffered data, wherein the buffered data was previously stored within the variable length buffer only if the variable length

buffer contained an amount of data less than a threshold based on the segment size.

7. The apparatus of claim 6 wherein the segment size is a Maximum Segment Size (MSS).

8. The apparatus of claim 7 wherein the threshold is approximately 3*MSS.

9. The apparatus of claim 6 wherein the buffered data comprises data from a Transmission Control Protocol/Internet Protocol (TCP/IP) session.

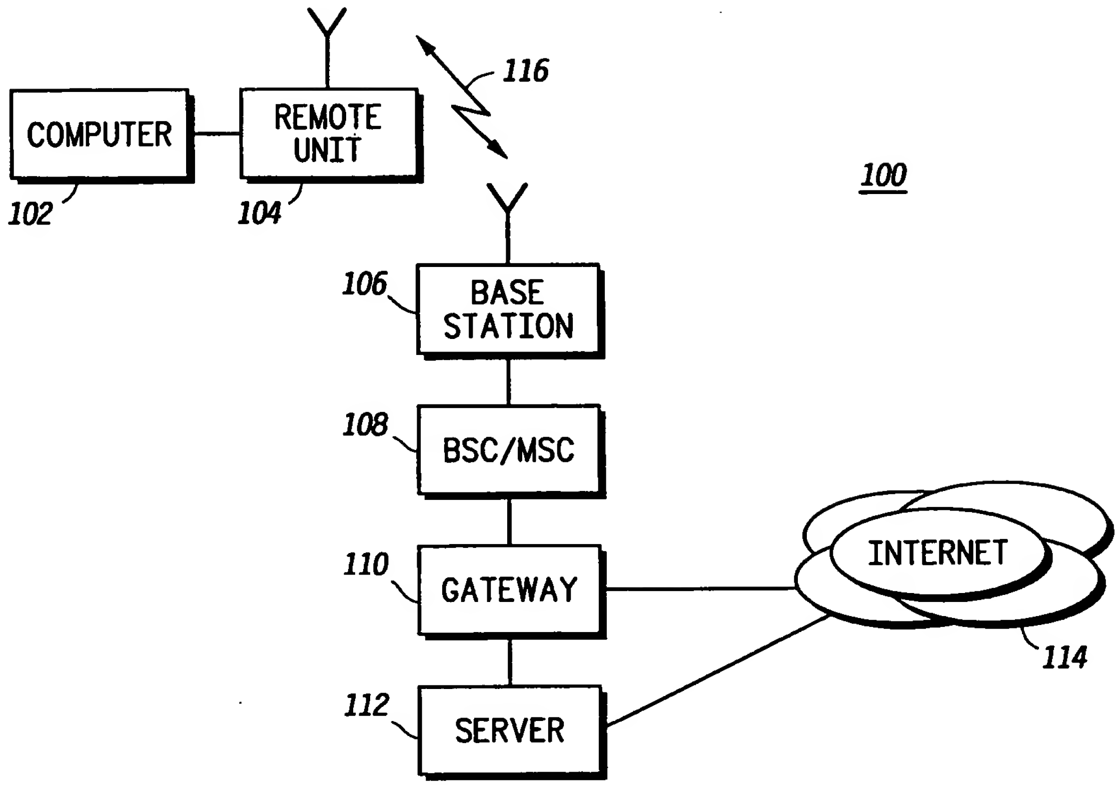

COMPUTER
REMOTE UNIT
116
100
102
104
106
BASE STATION
108
BSC/MSC
110
GATEWAY
INTERNET
114
112
SERVER

FIG.1

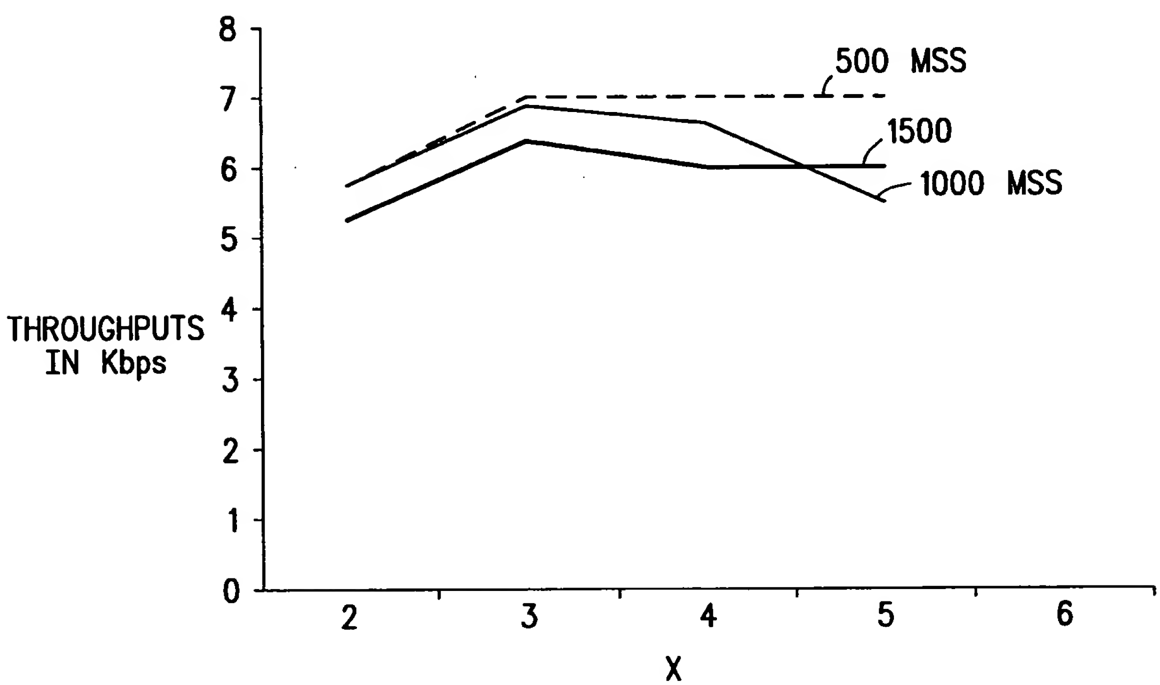

500 MSS
1500
1000 MSS
THROUGHPUTS IN Kbps
0
1
2
3
4
5
6
7
8
2
3
4
5
6
X

FIG.2

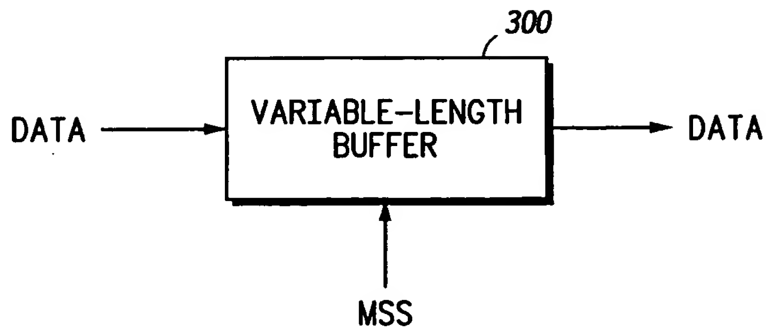

300
DATA
VARIABLE-LENGTH BUFFER
DATA
MSS

FIG.3

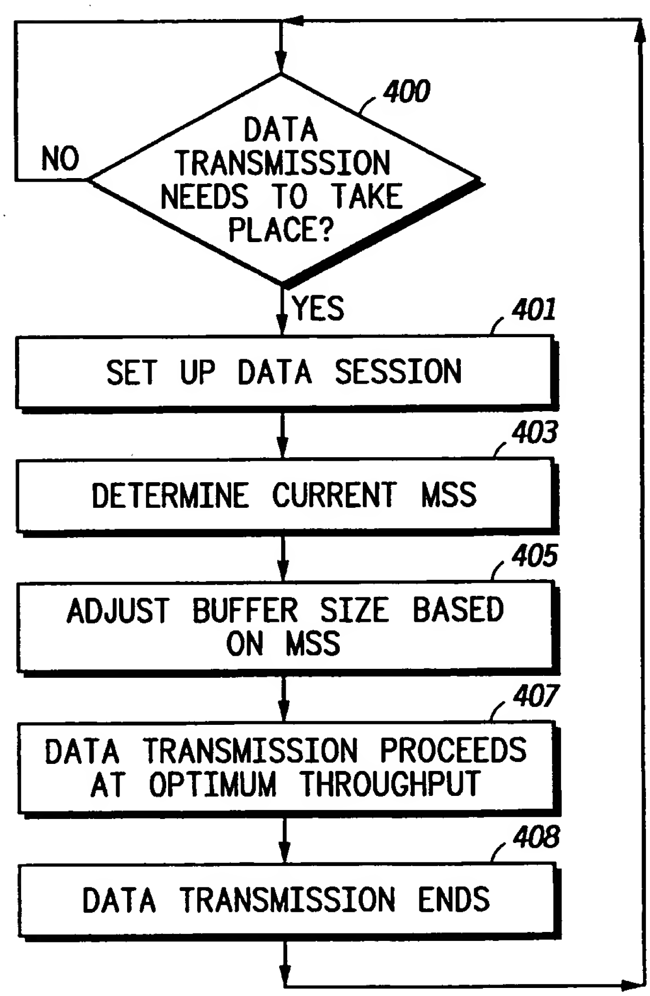

DATA TRANSMISSION NEEDS TO TAKE PLACE?
400
NO
YES
401
SET UP DATA SESSION
403
DETERMINE CURRENT MSS
405
ADJUST BUFFER SIZE BASED ON MSS
407
DATA TRANSMISSION PROCEEDS AT OPTIMUM THROUGHPUT
408
DATA TRANSMISSION ENDS

FIG.4

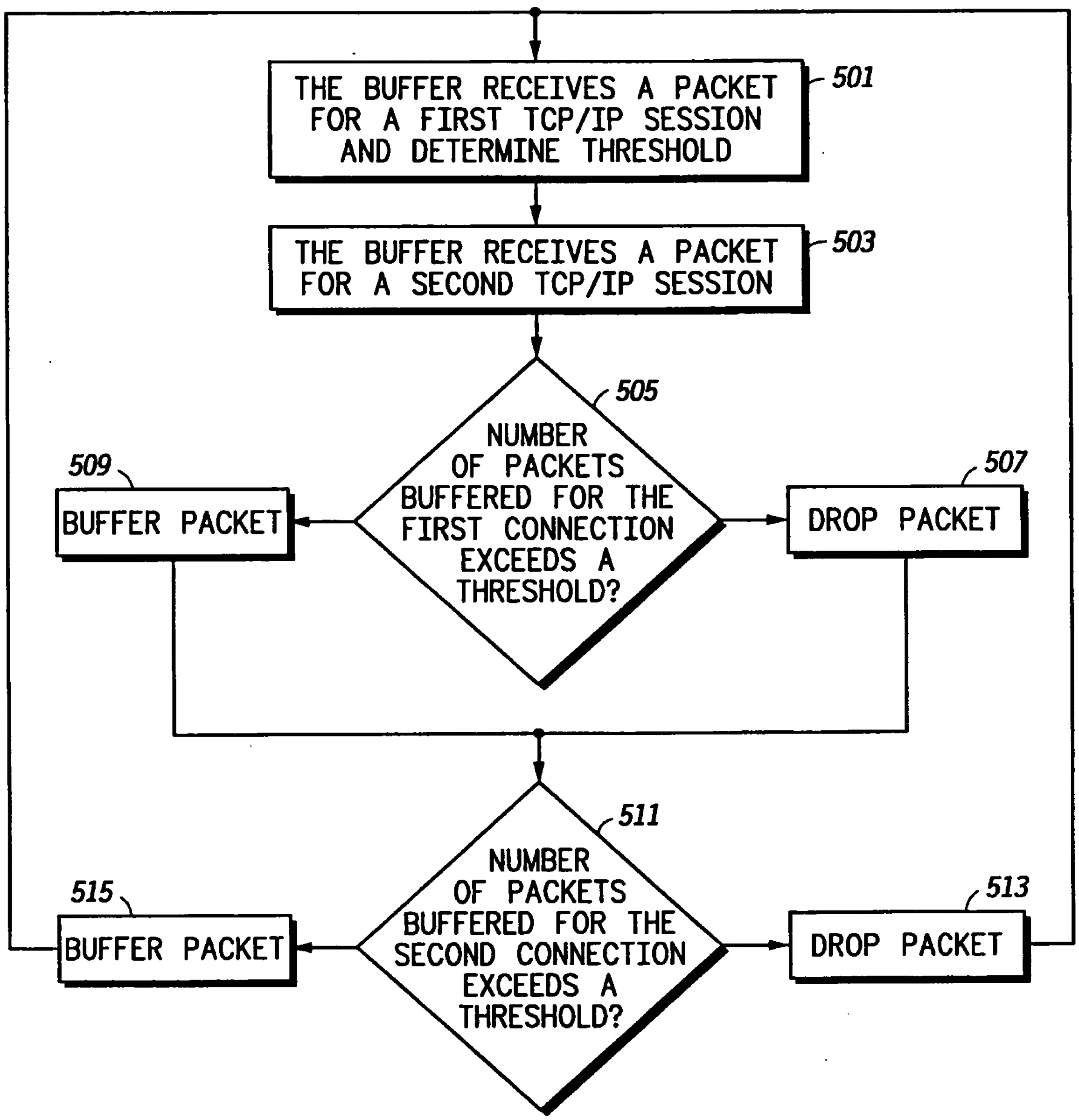

THE BUFFER RECEIVES A PACKET FOR A FIRST TCP/IP SESSION AND DETERMINE THRESHOLD
501
THE BUFFER RECEIVES A PACKET FOR A SECOND TCP/IP SESSION
503
NUMBER OF PACKETS BUFFERED FOR THE FIRST CONNECTION EXCEEDS A THRESHOLD?
505
509
BUFFER PACKET
507
DROP PACKET
NUMBER OF PACKETS BUFFERED FOR THE SECOND CONNECTION EXCEEDS A THRESHOLD?
511
515
BUFFER PACKET
513
DROP PACKET

FIG.5

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US00/10631

A. CLASSIFICATION OF SUBJECT MATTER
IPC(7) :H04L 12/54
US CL : 370/401, 428
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
U.S. : 370/328,389, 401

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EAST

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 5,793,762 A (PENNERS ET AL.) 11 AUGUST, 1998.	1-9
A,P	US 5,910,946 A (CSAPO) 08 JUNE 1999.	1-9

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier document published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
05 JUNE 2000	10 JUL 2000

Name and mailing address of the ISA/US	Authorized officer
Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231	DOUG OLMS
Facsimile No. (703) 305-3230	Telephone No. (703) 305-4703